УДК 519.7: 621.8

Р.М. Трохимчук
Київський національний університет імені Тараса Шевченка, Україна
пр. Академіка Глушкова, 4д, Київ, 03680

РЕЗУЛЬТАТИ ТЕСТУВАННЯ, ДОСЛІДЖЕННЯ ТА АНАЛІЗУ ОСНОВНИХ АЛГОРИТМІВ КЛАСТЕРИЗАЦІЇ НАБОРІВ ЧИСЛОВИХ ДАНИХ

R.M. Trokhymchuk
Kyiv National Taras Shevchenko University, Ukraine
4d, Academician Hlushkov Ave, Kyiv, 03680

RESULTS OF TESTING, RESEARCH AND ANALYSIS OF THE BASIC CLUSTERING ALGORITHMS OF NUMERICAL DATA SETS

Ця робота присвячена тестуванню, дослідженню та порівняльному аналізу найбільш відомих і широко використовуваних на практиці методів і алгоритмів кластеризації наборів числових даних. Для оцінки результатів розв’язання задачі кластеризації за допомогою візуалізації наборів даних на всіх етапах реалізації досліджуваних алгоритмів було застосовано метод багатовимірного шкалювання. Усі алгоритми були перевірені на штучних і реальних наборах даних. Для кожного з досліджених алгоритмів було сформульовано основні характеристики у вигляді їхніх відносних переваг і недоліків. На підставі результатів тестування сформульовано висновки і рекомендації щодо використання цих алгоритмів.

Ключові слова: інтелектуальний аналіз даних, кластерний аналіз, тестування алгоритмів, візуалізація даних, багатовимірне шкалювання

This work is devoted to the testing, research and comparative analysis of the most well-known and widely used methods and algorithms for clustering numerical data sets. Multidimensional scaling was applied to evaluate the results of solving the clustering problem by visualizing datasets at all stages of the implementation of the studied algorithms. All algorithms were tested for artificial and real data sets. As a result, for each of the investigated algorithms, the main characteristics were formulated in the form of their relative strengths and weaknesses. Based on the test results, conclusions and recommendations for using these algorithms are formulated.

Keywords: Data Mining, cluster analysis, algorithm testing, data visualization, multidimensional scaling (MDS)

Вступ

Ця робота містить результати тестування, дослідження і порівняльного аналізу найбільш відомих і широко використовуваних на практиці методів і алгоритмів кластеризації числових даних.

Кластеризація (інші назви: кластерний аналіз; класифікація, розпізнавання образів або навчання без вчителя; таксономія та ін.) широко і ефективно використовується в системах інтелектуального аналізу даних. Завданням інтелектуального аналізу даних є пошук у великих наборах даних прихованих важливих і корисних закономірностей, які дають змогу отримати нові знання про досліджувані дані. На сьогодні синонімами терміна «інтелектуальний аналіз даних» є видобування даних (Data Mining) і виявлення знань (Knowledge Discovery) [1,2].

Останнім часом особливий інтерес до методів інтелектуального аналізу даних виник у зв’язку з широким розповсюдженням і розвитком засобів збору і зберігання даних, які дають можливість накопичувати великі (величезні) обсяги інформації. Для фахівців з різних областей людської діяльності виникла проблема обробки та аналізу зібраних даних, перетворення їх у знання.

Популярні класичні математико-статистичні методи застосовні й ефективні для такого рівня далеко не у всіх ситуаціях. Для використання цих методів необхідно, як правило, мати попередні відомості (такі, наприклад, як: незалежність, однорідність, випадковість, вид розподілу тощо) про шукані закономірності та мати достатню кваліфікацію в галузі математичної статистики.

© Р.М. Трохимчук

У такій ситуації методи інтелектуального аналізу даних (які, крім іншого, безумовно, включають у себе і математико-статистичні методи) набувають особливої актуальності. Їхня основна особливість полягає у встановленні наявності і описі закономірностей у наборах даних, тоді як традиційні математико-статистичні методи орієнтовані головним чином на визначення або оцінку параметрів передбачуваних закономірностей.

Серед методів інтелектуального аналізу даних особливе місце займають класифікація та кластеризація. Класифікація, виходячи з відомого заздалегідь групування даних на підмножини (класи), встановлює закономірності, за якими дані групуються саме таким чином, і дає можливість у подальшому класифікувати (розпізнавати) нові невідомі об’єкти. Кластеризація ж, ґрунтуючись на певному відношенні схожості (подібності, близькості) елементів набору даних, формує підмножини (кластери), в які групуються вхідні дані.

Кластеризація (навчання без вчителя) істотно відрізняється від класифікації (навчання з учителем) тим, що невідомими є як приналежність окремих об’єктів початкової вибірки до певних класів (кластерів), так і число таких класів.

Можна виділити такі основні цілі задачі кластеризації:

- *Розуміння даних*. Розбиття заданої множини об’єктів (початкової вибірки) на групи подібні між собою дає змогу визначити структуру цієї множини. Це, у свою чергу, дає можливість спростити по-дальшу обробку даних.
- *Стиснення даних*. Скорочення обсягу збережених даних шляхом формування репрезентативної вибірки, тобто виділення і збереження найбільш типових представників у кожному кластері (наприклад, центрів кластерів, кількох центральних елементів з кожного кластера тощо).
- *Виявлення новизни*. Виділення нетипових (особливих) об’єктів, які не входять до жодного зі знайдених кластерів.
- *Побудова формальної (математичної) моделі* для опису механізмів і процесів породження аналізованих даних, отримання можливості екстраполяції (передбачення поведінки) таких процесів.
- *Розпізнавання образів*. Набір даних, для якого вирішена задача кластеризації, може в подальшому стати основою (як навчена початкова вибірка) для задачі класифікації.
- *Тестування даних*. У багатьох випадках, коли виникають сумніви у «кваліфікації вчителя» в задачі класифікації, буває корисно навчену початкову вибірку піддати кластеризації, щоб переконатися в достовірності заданої вчителем структури цієї вибірки.

Постановка завдання

Задача кластеризації може бути формалізована у такий спосіб. Задано початкову вибірку (множину об’єктів) X і функцію ρ близькості (подібності) між цими об’єктами. Потрібно розбити вибірку X на непересічні підмножини, які називаються кластерами, так, щоб кожен кластер складався з об’єктів, близьких згідно з заданою функцією ρ, а об’єкти різних кластерів істотно відрізнялися.

Алгоритм кластеризації – це процедура визначення функції, яка кожному об’єкту з X приписує мітку (номер) відповідного кластера. Множина міток рідко буває відома заздалегідь, тому часто в задачу кластеризації входить також визначення оптимального числа кластерів, з точки зору того чи іншого критерія якості кластеризації.

Таким чином, задачу кластеризації можна сформулювати як задачу дискретної оптимізації: необхідно так приписати номера кластерів об’єктам початкової вибірки, щоб значення певного функціоналу якості стало оптимальним. Існує багато різновидів функціоналів якості кластеризації, але немає «найкращого» серед них.

Рішення задачі кластеризації принципово неоднозначно з ряду причин. По-перше, не існує універсального найкращого критерія якості кластеризації. По-друге, число кластерів, як правило, невідомо заздалегідь і встановлюється відповідно до деяких

суб'єктивних критеріїв. По-третє, результат кластеризації істотно залежить від функції близькості ρ, вибір якої, як правило, також суб'єктивний. Нарешті, результат кластеризації залежить від застосовуваного для її вирішення алгоритму. Вище перераховано лише основні причини неоднозначності.

Створення алгоритму кластеризації, що успішно працює в усіх ситуаціях, є задачею нереальною і безперспективною. Також слід враховувати, що на сьогодні не існує формальних способів адекватного вибору конкретного алгоритму (або алгоритмів) кластеризації для заданих наборів даних. Таким чином, задача кластеризації у більшості випадків є суто евристичною.

Оскільки задача кластеризації може бути вирішена різними способами, то для якісного і швидкого її розв'язання в конкретній ситуації необхідно мати методики вибору найбільш адекватних з можливих процедур. У зв'язку з цим особливої актуальності набуває знання особливостей, основних характеристик, переваг і недоліків різних доступних методів і алгоритмів кластеризації. Таке знання істотно полегшить вибір найкращого рішення і дозволить отримувати найбільш достовірний результат.

У цій роботі розглянуто й проаналізовано найбільш відомі і широко використовувані на практиці алгоритми кластеризації, призначені для обробки числових даних. Здійснено тестування цих алгоритмів на штучних і реальних даних. За результатами тестування сформульовано висновки і рекомендації.

Методика дослідження

Багаторазові спроби класифікації методів кластерного аналізу призводять до десятків, а то і сотень різноманітних класів алгоритмів кластеризації. Таке розмаїття породжується великою кількістю можливих способів обчислення близькості між окремими об'єктами вибірки, не меншою кількістю методів обчислення близькості між окремими кластерами, різноманітними оцінками (критеріями) оптимальності кінцевої кластерної структури та ін. Загальноприйнятої класифікації методів кластеризації не існує, але можна виділити ряд груп підходів (деякі методи можна віднести відразу до кількох груп).

Найбільшого поширення набули дві групи алгоритмів кластерного аналізу: ієрархічні і неієрархічні (ітеративні) методи [3,4].

Основними методами ієрархічного кластерного аналізу є агломеративні методи (AGNES – Agglomerative Nesting) одиночного (ближнього сусіда), повного і середнього зв'язку, дівізимні методи (DIANA – Divisive Analysis) BIRCH, MST, метод Варда. Серед неієрархічних методів слід виділити алгоритм k-середніх, PAM (Partitioning Around Medoids) – k-means + k-medoids, MeanShift, EM-алгоритм (Expectation-Maximization), алгоритми DBSCAN і FOREL.

Вибираючи між ієрархічними і неієрархічними методами, слід звернути увагу на такі моменти. Неієрархічні методи є більш стійкими до викидів, невдалого вибору функції близькості, використання неістотних (фіктивних) змінних серед координат наборів даних та ін. Але використовуючи ці методи, дослідник повинен обирати початкову (стартову) точку, підсумкову кількість кластерів, умову зупинки алгоритму та інші параметри. Все це істотно відбивається на ефективності і часі роботи ітеративних алгоритмів.

У даній роботі для аналізу було обрано такі алгоритми:

- Спектральна кластеризація (Spectral Clustering);
- Ієрархічна (агломераційна) кластеризація з одиночним і повним зв'язком (Agglomerative clustering with single and complete linkage);
- Середнє зміщення (Mean Shift);
- BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies);
- k-середніх (k-Means і MiniBatch-k-Means);
- CURE (Clustering Using REpresentatives);
- k-Medoids;
- DBSCAN (Density-Based Spatial Clustering of Applications with Noise);
- EM (Expectations-Maximization);
- FOREL (Formal Element).

© Р.М. Трохимчук

Відповідні програми були взяті з різних відкритих джерел. У процесі апробації деякі з цих програм модифікувалися для адаптації до конкретної ситуації.

Було розроблено дві групи тестів для аналізу і порівняння алгоритмів кластеризації. Перша група складається з модельних (штучних) наборів точок, для яких результат можливої кластеризації заздалегідь відомий. З метою перевірки досліджуваних алгоритмів на стабільність для цієї групи застосовувалися методи накладання різних шумів на початкові ідеальні дані. У другій групі тестів точки початкової вибірки вибираються випадково за допомогою різних датчиків випадкових чисел із застосуванням різних законів розподілу.

Для оцінки результатів розв'язання задачі кластеризації істотне значення мають прості й зручні у використанні засоби візуалізації та здійснена з їхньою допомогою експертна оцінка достовірності цих результатів. Візуалізація даних є важливою частиною якісної системи аналізу даних. Візуалізацію бажано запроваджувати як для початкового набору даних, так і для аналізу проміжних та, особливо, остаточних результатів кластеризації. Візуалізація наборів даних є непростою проблемою, особливо якщо початкова вибірка велика, а простір об'єктів істотно багатовимірний.

Для візуалізації початкової вибірки і результатів кластеризації використовують метод багатовимірного шкалювання (MultiDimensional Scaling, MDS) [5,6], який дає можливість відобразити початкову вибірку й рішення задачі кластеризації у вигляді множини точок у просторі меншої розмірності. Наприклад, отримати тривимірне або навіть плоске відображення для наборів даних. Таке подання в цілому відображає основні структурні особливості заданої багатовимірної вибірки, зокрема, її кластерну структуру. Тому три- або двовимірне шкалювання часто використовують для аналізу і розуміння як початкових даних, так і результатів рішення задачі кластеризації.

Слід зазначити, що в роботі [7] метод багатовимірного шкалювання використовувався як для візуалізації наборів даних і результатів кластеризації, так і для зменшення розмірності простору ознак. Точніше, було здійснено ранжування елементів множини ознак з метою вибору найбільш інформативних та істотних ознак. Безумовно, такий підхід до реалізації кластеризації наборів даних слід рекомендувати як надзвичайно перспективний і такий, що заслуговує найпильнішої уваги.

Саме метод багатовимірного шкалювання використовувався в даній роботі для аналізу, оцінки та порівняння різних алгоритмів кластеризації.

Нарешті, важливим етапом у вирішенні задачі кластеризації є змістовна інтерпретація результатів кластеризації. Зокрема, опис отриманих кластерів мовою предметної області. Як правило, цю частину рішення доцільно доручати кваліфікованим фахівцям у даній області. При цьому ефективним інструментом для здійснення адекватної інтерпретації результатів кластеризації є зручна, наочна й зрозуміла для звичайних користувачів візуалізація підсумкового розподілу даних на кластери.

Результати дослідження

Ідеальним результатом порівняння алгоритмів кластеризації були б різні кількісні показники для оцінки тих чи інших характеристик і особливостей цих алгоритмів. Отримати такі показники можливо (і нескладно) в кожній конкретній ситуації. Однак значення і цінність таких даних невеликі, тому що значення цих показників істотно залежать від особливостей того початкового набору даних, для якого застосовувався аналізований алгоритм кластеризації. Найоб'єктивніший кількісний порівняльний аналіз можна було б отримати лише використовуючи певні стандартні або канонічні (наприклад, за типом, структурою і розміром) набори даних, яких на сьогоднішній день не існує.

Зокрема, при аналізі кожного з перерахованих вище алгоритмів безумовно визначалися традиційні кількісні параметри їхньої реалізації – час роботи і обсяг споживаної пам'яті. Однак продуктивність сучасної об-

© Р.М. Трохимчук

числювальної техніки така, що значення цих параметрів не є аж так істотними для розв'язання більшості реальних задач кластеризації. Тому в цьому огляді дані параметри відображені лише опосередковано.

У результаті для кожного з досліджених алгоритмів було сформульовано основні характеристики у вигляді їхніх відносних переваг і недоліків.

Спектральна кластеризація

Переваги: добре працює для невеликої кількості кластерів.

Недоліки: повільний; вимагає зазначення кількості кластерів; не рекомендується для випадку великого числа кластерів.

Агломераційна кластеризація з одиночним зв'язком

Переваги: добре працює для невеликої кількості кластерів (до 10000 точок); не вимагає попереднього задання кількості кластерів; може розділити дані на будь-яку кількість кластерів; велика гнучкість при зміні значень параметрів і обмежень; хороші результати для кластерів складної структури; досить добре масштабується.

Недоліки: алгоритм досить повільний; має тенденцію створювати довгі тонкі кластери, в яких сусідні елементи одного кластера близькі, в той час як елементи на протилежних кінцях кластера можуть бути набагато далі один від одного, ніж два елементи різних кластерів; погано працює з неопуклими кластерами.

Агломераційна кластеризація з повним зв'язком

Переваги: практично ті самі, що й для одиночного зв'язку.

Недоліки: порядок обробки даних впливає на кінцевий результат; чутливий до викидів («шуму»); висока обчислювальна складність.

Середнє зміщення (Mean Shift)

Переваги: висока швидкість реалізації; автоматично встановлює кількість кластерів.

Недоліки: не є добре масштабованим, так як вимагає під час виконання постійного пошуку найближчого сусіда; відсутність обґрунтування та гарантій збіжності алгоритму до оптимального рішення; припиняє пошук рішення, коли зміна в центроїдах є малою.

BIRCH

Переваги: двоступенева кластеризація; можливість кластеризації великих об'ємів даних; обмежений обсяг пам'яті; може працювати за один прохід, але дає змогу поліпшити якість рішення за допомогою кількох додаткових запусків; успішно застосовується для неоднорідних за розмірами та формами кластерів; добре масштабується для порівняно невеликого числа кластерів; успішно справляється з ситуацією наявності «шуму» у початкових даних.

Недоліки: обробляє дані тільки числових типів; вимагає задання порогових значень; добре виділяє тільки кластери опуклої або сферичної форми; погано масштабується для великих наборів даних.

K-Means

Переваги: простота налаштування й використання; хороша швидкість реалізації; зрозумілість і прозорість алгоритму; дає хороші результати для опуклих даних; добре масштабується.

Недоліки: алгоритм занадто чутливий до викидів; повільна робота для великих наборів даних; необхідність задавати кількість кластерів; неможливість застосування алгоритму для даних, де кластери перетинаються; рандомізований, що означає можливість отримання різних результатів при кожному його запуску; відсутність гарантії отримання оптимального рішення; погано працює, коли розміри і форми кластерів неоднорідні.

MiniBatch-K-Means є модифікацією алгоритму K-Means, яка дає змогу істотно підвищити ефективність рішення в порівнянні з оригінальним алгоритмом.

Переваги: простота і зрозумілість алгоритму та його використання; більш висока швидкість реалізації порівняно з методом K-Means; можливість клас-

© Р.М. Трохимчук

теризації великих об’ємів даних; хороший баланс якості рішення та часу обчислень.

Недоліки: результат істотно залежить від ініціалізації центроїдів, тому бажаною є багаторазова реалізація процедури.

CURE. Алгоритм кластеризації CURE є ще однією модифікацією методу K-Means з метою усунення ситуації, коли розміри і форми кластерів неоднорідні.

Переваги: якісно виконує кластеризацію навіть при наявності викидів; виділяє кластери складної форми і різних розмірів; не вимагає великих затрат пам’яті.

Недоліки: необхідність у заданні порогових значень і кількості кластерів; погано застосовний для великих наборів даних з огляду на велику часову складність.

K-medoids

Переваги: простота використання; висока швидкість реалізації; зрозумілість і прозорість алгоритму; алгоритм менш чутливий до викидів у порівнянні з K-Means.

Недоліки: необхідно задавати кількість кластерів; повільна робота на великих наборах даних.

DBSCAN

Переваги: найкраще працює на щільних кластерах; гарантує оптимальні рішення при правильному виборі параметрів; нечутливість до викидів; здатність виділяти кластери довільної форми; не вимагає задання кількості кластерів і автоматично визначає це число.

Недоліки: алгоритм не дуже надійний, тому що дуже чутливий до зміни параметрів; рандомізований; досить складний у налаштуванні, бо непросто знайти адекватні значення параметрів; час реалізації досить великий; для отримання найкращого результату слід запускати його кілька разів з різними комбінаціями параметрів.

EM (Expectations-Maximization)

Переваги: стійкість до шумів і викидів; можливість розбиття початкового набору даних на заздалегідь задану кількість кластерів; добре математично обґрунтований; можливість його застосування для даних, у яких кластери перетинаються; лінійна залежність складності реалізації від розміру набору даних; швидка збіжність алгоритму при вдалому виборі початкових умов.

Недоліки: бажано, щоб всі параметри даних були нормально розподілені; відсутність гарантії отримання оптимального рішення, оскільки алгоритм може зупинитися в локальному мінімумі й дати квазіоптимальний розв’язок.

FOREL

Переваги: точність мінімізації функціоналу якості (при вдалому підборі основного параметра R); наочність візуалізації результатів кластеризації; гарантована (математично обґрунтована) збіжність алгоритму; можливість оперативно втручатися в роботу алгоритму, здійснюючи корекцію центрів кластерів; можливість підрахунку проміжних значень різних функціоналів якості; можливість перевірки гіпотез схожості і компактності в процесі роботи алгоритму; не вимагає задання кількості кластерів.

Недоліки: відносно низька продуктивність (швидкість реалізації); незадовільні результати для кластерів складної форми; нестійкість алгоритму (залежність від вибору початкової точки); необхідність апріорних знань про основні характеристики кластерів.

Висновки і рекомендації

Одним з центральних критеріїв якості рішення задачі кластеризації є характеристика, яку можна назвати стабільністю рішення. Отримане рішення задачі кластеризації можна вважати стабільним, якщо цей результат зберігається при зміні методів кластеризації.

Важко розраховувати на повний збіг результатів при застосуванні різних процедур кластеризації. Тому на практиці рішення вважається стабільним, якщо при порівнянні групи збігаються більше, ніж на 70%. Тут діє просте емпіричне правило – стійка типо-

логія зберігається при зміні методів кластеризації.

Наслідком перевірки отриманої кластеризації на стабільність є висновок про достовірність (об'єктивність) розв'язання задачі. На сьогодні перевірити достовірність (адекватність) отриманого рішення іншими методами, не надається можливим. Саме цей метод багаторазового застосування кількох алгоритмів кластеризації до заданого набору даних слід рекомендувати на практиці для отримання максимально достовірного рішення.

Безумовно, слід рекомендувати також використання різних процедур візуалізації даних на всіх етапах виконання задачі кластеризації. Це доцільно робити спочатку для дослідження особливостей початкового набору даних. І особливо важливою є ця процедура для аналізу підсумкових результатів реалізації алгоритмів кластеризації. На сьогодні крім вищезгаданого методу багатовимірного шкалювання MDS [5,6,7] існують також інші доступні й досить прості у використанні процедури візуалізації даних.

Література

1. Negnivitsky, M. (2002). *Artificial Intelligence A Guide to Intelligent Systems*. Harlow: Addison-Wesley, Pearson Education Limited.
2. Data Mining, Web Mining, Text Mining, and Knowledge Discovery. URL: http://kdnuggets.com.
3. Tan P.-N., Steinbach M., Karpatne A., Kumar V. (2017). *Introduction to Data Mining (What's New in Computer Science)*. Addison-Wesley.
4. Бериков, В.С., Лбов, Г.С. (2008). Современные тенденции в кластерном анализе. Всероссийский конкурсный отбор обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы».
5. Толстова, Ю.Н. (2006). *Основы многомерного шкалирования*. Москва: КДУ.
6. Кулаичев, А.П. (2006). *Методы и средства комплексного анализа данных*. Москва: ФОРУМ: ИНФРА-М.
7. Krak, I.V., Kudin, G.I., Kulyas, A.I. (2019). Multidimensional Scaling by Means of Pseudoinverse Operations. *Cybernetics and Systems Analysis*. Vol. 55, Iss. 1. P. 30-38.

References

1. Negnivitsky, M. (2002). *Artificial Intelligence A Guide to Intelligent Systems*. Harlow: Addison-Wesley, Pearson Education Limited.
2. Data Mining, Web Mining, Text Mining, and Knowledge Discovery. URL: http://kdnuggets.com.
3. Tan P.-N., Steinbach M., Karpatne A., Kumar V. (2017). *Introduction to Data Mining (What's New in Computer Science)*. Addison-Wesley.
4. Berykov, V.S., Lbov, H.S. (2008). Sovremennye tendentsyy v klasternom analyze. Vserossyyskyy konkursnyy otbor obzorno-analytycheskykh statey po pryorytetnomu napravlenyyu «Ynformatsyonno-telekommunykatsyonnye systemy».
5. Tolstova, Yu.N. (2006). Osnovy mnohomernoho shkalyrovanyya. Moskva: KDU.
6. Kulaychev, A.P. (2006). Metody y sredstva kompleksnoho analyza dannykh. Moskva: FORUM: YNFRA-M.
7. Krak, I.V., Kudin, G.I., Kulyas, A.I. (2019). Multidimensional Scaling by Means of Pseudoinverse Operations. *Cybernetics and Systems Analysis*. Vol. 55, Iss. 1. P. 30-38.

RESUME

R.M. Trokhymchuk

Results of testing, research and analysis of the basic clustering algorithms of numerical data sets

This work is devoted to the testing, research and comparative analysis of the most well-known and widely used methods and algorithms (hierarchical and non-hierarchical or iterative) for clustering numerical data sets.

Two groups of tests were developed for analyzing and comparing clustering algorithms. The first group consists of model (artificial) sets of points for which the result of possible clustering is predefined. In order to test the applied algorithms for stability for this group, various methods of superimposing various noises on the original ideal data were used. In the second group of tests, the points of the initial sample are chosen randomly using various random number sensors using different distribution laws.

Multidimensional scaling was applied to evaluate the results of solving the clustering problem by visualizing datasets at all stages of the implementation of the studied algorithms.

For each of the algorithms studied the main characteristics are formulated in the form of their relative advantages and disadvantages. According to the test results, conclusions and recommendations for using these algorithms are formulated.

Надійшла до редакції 12.06.2019

© Р.М. Трохимчук